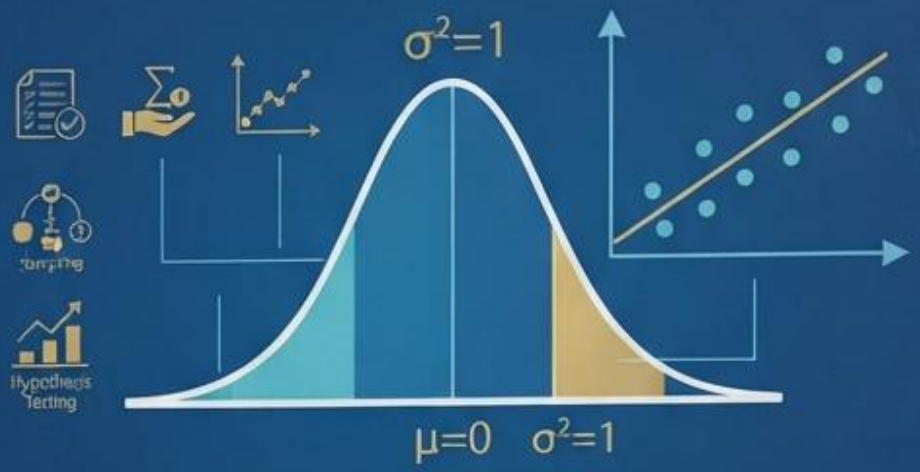
MAR 121
APPLIED STATISTICS



အခန်း ၁: အညွှန်းကိန်းများ (Index Numbers)

အခန်း ၂: ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution)

အခန်း ၃: နမူနာကောက်ယူခြင်းနှင့် ခန့်မှန်းခြေ (Sampling and Estimation)

အခန်း ၄: ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing)

အခန်း ၅: ဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်း (Regression Analysis)

အခန်း ၆: အမှတ်ခန့်မှန်းခြေ (Point Estimation)

အခန်း ၇: ကြားကာလခန့်မှန်းခြေ (Interval Estimation)

အခန်း ၈: ယူဆချက်စစ်ဆေးခြင်း (Tests of Statistical Hypotheses)

အခန်း ၉: နောက်ထပ်စစ်ဆေးမှုများ (More Tests)

ACADEMIC TEXTBOOK | DEPARTMENT OF STATISTICS | UNIVERSITY OF COMPUTER STUDIES

အခန်း (၁) - အညွှန်းကိန်းများ (Index Numbers)

ဤအခန်းတွင် အညွှန်းကိန်း (Index Number) ဟူသည့် အကြောင်းအရာကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၁.၁ နိဒါန်း (Introduction)

အညွှန်းကိန်း (Index Number) ဆိုသည်မှာ အချိန်ကာလတစ်ခုအတွင်း ဆက်စပ်နေသော ပစ္စည်းများ (သို့) ကိန်းရှင်များ၏ ပြောင်းလဲမှုကို တိုင်းတာသည့် ကိရိယာတစ်ခုဖြစ်သည်။ ဤပြောင်းလဲမှုများသည် ကုန်ပစ္စည်းဈေးနှုန်း၊ ထုတ်လုပ်မှုပမာဏ၊ ရောင်းအား သို့မဟုတ် စားသုံးမှုပမာဏ စသည်တို့နှင့် သက်ဆိုင်နိုင်သည်။ ထို့အပြင် ဉာဏ်ရည်၊ အလှအပ သို့မဟုတ် ထိရောက်မှုကဲ့သို့သော စိတ်ကူးယဉ်ဆန်သည့် အရာများကိုလည်း တိုင်းတာနိုင်သည်။

အညွှန်းကိန်းကို အောက်ပါတို့အတွက် အသုံးပြုနိုင်သည်-

- ကာလအလိုက် နှိုင်းယှဉ်ခြင်း (ဥပမာ - ယခုနှစ်နှင့် လွန်ခဲ့သော ၅ နှစ်က ကုန်ဈေးနှုန်း)။
- နေရာဒေသအလိုက် နှိုင်းယှဉ်ခြင်း (ဥပမာ - ရန်ကုန်နှင့် မန္တလေးရှိ နေထိုင်စရိတ်)။
- အုပ်စုအလိုက် နှိုင်းယှဉ်ခြင်း (ဥပမာ - ကျောင်းတစ်ကျောင်း၏ ထိရောက်မှု)။

ဥပမာအားဖြင့်၊ ဈေးနှုန်းအညွှန်းကိန်း (Price Index) ဆိုသည်မှာ သတ်မှတ်ထားသော ကုန်ပစ္စည်းအုပ်စုတစ်ခု၏ ဈေးနှုန်းပြောင်းလဲမှုကို ကိန်းဂဏန်းတစ်ခုတည်းဖြင့် ဖော်ပြခြင်းဖြစ်သည်။

၁.၂ အညွှန်းကိန်းတည်ဆောက်ခြင်း (Construction of an Index Number)

အညွှန်းကိန်းတစ်ခုကို တည်ဆောက်ရာတွင် အခြေခံကျသော နည်းလမ်းမှာ ပမာဏ (သို့) ဈေးနှုန်းတို့၏ ဆွေမျိုးပြောင်းလဲမှု (relative change) ကို ပျမ်းမျှတွက်ချက်ခြင်းဖြစ်သည်။

ဥပမာ-

လွန်ခဲ့သောနှစ်က ဂျုံထွက်ရှိမှုသည် ၁၁၀% ရှိပြီး ချည်ထွက်ရှိမှုသည် ၁၀၆% ဖြစ်သည် ဆိုပါစို့။ ဤရာခိုင်နှုန်းနှစ်ခုကို ပျမ်းမျှတွက်လျှင် စုစုပေါင်းထွက်ရှိမှုသည် ယခင်နှစ်၏ ၁၀၈% ရှိသည်ဟု

ဆိုနိုင်သည်။ သို့သော် ဤသို့တွက်ချက်ရာတွင် ပစ္စည်းနှစ်မျိုးလုံး၏ အရေးပါမှု (importance) တူညီသည်ဟု ယူဆထားသည်။

အကယ်၍ ချည်ထွက်ရှိမှုသည် ဂျုံထွက်ရှိမှုထက် ၆ ဆ ပိုအရေးပါပါက၊ ရာခိုင်နှုန်းများကို အလေးချိန်သတ်မှတ်၍ တွက်ချက်ရမည် (ဥပမာ - ချည်အတွက် ၆၊ ဂျုံအတွက် ၁)။ ဤသို့ အလေးချိန်သတ်မှတ်တွက်ချက်သော ပျမ်းမျဆွေမျိုးကို **အညွှန်းကိန်း (Index Number)** ဟုခေါ်သည်။

အခြေခံကာလ (Base Period)

အညွှန်းကိန်းအားလုံး၏ နောက်ထပ်ထူးခြားချက်တစ်ခုမှာ ၎င်းတို့ကို သတ်မှတ်ထားသော နေရာ၊ အချိန် သို့မဟုတ် ကာလတစ်ခုနှင့် နှိုင်းယှဉ်ဖော်ပြခြင်းဖြစ်သည်။ ဤသတ်မှတ်ထားသော ကာလကို **အခြေခံကာလ (Base Period)** ဟုခေါ်သည်။

၁.၃ ရိုးရှင်းသော အညွှန်းကိန်း (Simple Index Numbers)

ရိုးရှင်းသော အညွှန်းကိန်းဆိုသည်မှာ ပစ္စည်းတစ်မျိုးတည်းကိုသာ အခြေခံ၍ တည်ဆောက်ထားသော အညွှန်းကိန်းဖြစ်သည်။

အသုံးအနှုန်းများ-

- **လက်ရှိကာလ (Current Period):** သင်သည် အညွှန်းကိန်းကို ရှာဖွေလိုသော ကာလ။
- **အခြေခံကာလ (Base Period):** လက်ရှိကာလ၏ ဈေးနှုန်းကို နှိုင်းယှဉ်လိုသော ကာလ။

အခြေခံကာလကို ရွေးချယ်ရာတွင် အောက်ပါအချက်များကို ဂရုပြုသင့်သည်-

၁။ ပုံမှန်ကာလတစ်ခုဖြစ်ရမည် (ဆိုလိုသည်မှာ ဈေးနှုန်းတည်ငြိမ်သောကာလ)။
၂။ အတိတ်ကာလကို ဝေးလွန်းအောင် မရွေးချယ်သင့်ပါ (စီးပွားရေးအပြောင်းအလဲများကြောင့် မကြာခဏ အခြေခံကာလကို ပြောင်းလဲလေ့ရှိသည်)။

၁.၃.၁ ဈေးနှုန်းဆွေမျိုး (Price Relatives)

ဈေးနှုန်းဆွေမျိုး (Price Relative) ဆိုသည်မှာ ပစ္စည်းတစ်မျိုး၏ လက်ရှိဈေးနှုန်းကို အခြေခံကာလ၏ ဈေးနှုန်းဖြင့် စားထားသော အချိုးဖြစ်သည်။

သင်္ကေတများ-

- p_0 = အခြေခံကာလ၏ ဈေးနှုန်း
- p_n = လက်ရှိကာလ၏ ဈေးနှုန်း

ဈေးနှုန်းဆွေမျိုး (Price Relative) = $\frac{p_n}{p_0}$

၁.၃.၂ ရိုးရှင်းသော ဈေးနှုန်းအညွှန်းကိန်း (Simple Price Index)

ရိုးရှင်းသော ဈေးနှုန်းအညွှန်းကိန်းဆိုသည်မှာ ဈေးနှုန်းဆွေမျိုးကို ရာခိုင်နှုန်းဖြင့် ဖော်ပြခြင်းဖြစ်သည်။

ရိုးရှင်းသော ဈေးနှုန်းအညွှန်းကိန်း (Simple Price Index) = ဈေးနှုန်းဆွေမျိုး × ၁၀၀

ဥပမာ ၁.၁

တစ်နှစ်အတွင်း ဆန်တစ်ပြည်လျှင် ဈေးနှုန်းသည် ၂၀၀၀ ကျပ်မှ ၂၅၀၀ ကျပ်သို့ တက်သွားပါက၊

- အခြေခံကာလ ဈေးနှုန်း (p_0) = ၂၀၀၀ ကျပ်
- လက်ရှိကာလ ဈေးနှုန်း (p_n) = ၂၅၀၀ ကျပ်
- ဈေးနှုန်းဆွေမျိုး = $\frac{2500}{2000} = 1.25$
- ရိုးရှင်းသော ဈေးနှုန်းအညွှန်းကိန်း = $1.25 \times 100 = 125$

ဆိုလိုသည်မှာ ဆန်ဈေးနှုန်းသည် ၂၅% တက်သွားသည်ဟု ဆိုလိုသည်။

၁.၄ ပေါင်းစပ်အညွှန်းကိန်း (Composite Index Number)

ပေါင်းစပ်အညွှန်းကိန်းဆိုသည်မှာ ပစ္စည်းတစ်မျိုးထက်ပိုသော အုပ်စုတစ်ခုလုံး၏ ဈေးနှုန်း သို့မဟုတ် ပမာဏပြောင်းလဲမှုကို တိုင်းတာခြင်းဖြစ်သည်။

၁.၄.၁ ရိုးရှင်းသော စုစုပေါင်းအညွှန်းကိန်း (Simple Aggregate Index)

ဤနည်းလမ်းတွင် မည်သည့်နှစ်၏ စုစုပေါင်းဈေးနှုန်းကိုမဆို အခြေခံနှစ်၏ စုစုပေါင်းဈေးနှုန်း၏ ရာခိုင်နှုန်းအဖြစ် ဖော်ပြသည်။

ရိုးရှင်းသော စုစုပေါင်းအညွှန်းကိန်း (Simple Aggregate Index) = $\frac{\sum p_n}{\sum p_0} \times 100$

- $\sum p_n$ = လက်ရှိကာလ၏ ဈေးနှုန်းများ ပေါင်းလဒ်
- $\sum p_0$ = အခြေခံကာလ၏ ဈေးနှုန်းများ ပေါင်းလဒ်

အားနည်းချက်များ-

၁။ ဈေးနှုန်းကြီးသော ပစ္စည်းတစ်ခုတည်းက အညွှန်းကိန်းကို လွှမ်းမိုးနိုင်သည်။
၂။ ပစ္စည်းတစ်ခုချင်းစီ၏ ရောင်းအားကို ထည့်သွင်းစဉ်းစားခြင်းမရှိပါ။
၃။ ပစ္စည်းများ၏ ယူနစ်ကို ပြောင်းလဲပါက အညွှန်းကိန်းတန်ဖိုး ပြောင်းလဲသွားနိုင်သည်။

ဤအားနည်းချက်များကြောင့် ရိုးရှင်းသော စုစုပေါင်းအညွှန်းကိန်းကို မကြာခဏ အသုံးမပြုကြပေ။

၁.၄.၂ ဆွေမျိုးဈေးနှုန်းများ၏ ပျမ်းမျှ (Average of Relative Prices)

ဤနည်းလမ်းသည် ရိုးရှင်းသော စုစုပေါင်းအညွှန်းကိန်း၏ အားနည်းချက်ကို ကျော်လွှားရန် ကြိုးစားထားသည်။ ၎င်းသည် ပစ္စည်းတစ်ခုချင်းစီ၏ ရိုးရှင်းသော ဈေးနှုန်းအညွှန်းကိန်းများ၏ ပျမ်းမျှဖြစ်သည်။

ဆွေမျိုးဈေးနှုန်းများ၏ ပျမ်းမျှ (Average of Relative Prices) = $\frac{\sum\left(\frac{p_n}{p_0} \times 100\right)}{k}$

- k = ပစ္စည်းအရေအတွက်

အားနည်းချက်-
ဤနည်းလမ်းသည်လည်း ပစ္စည်းတစ်ခုချင်းစီ၏ ရောင်းအားကို ထည့်သွင်းစဉ်းစားခြင်း မရှိသေးပါ။ ပစ္စည်းတိုင်း၏ ရောင်းအား တူညီသည်ဟု ယူဆထားသည်။

၁.၅ အလေးချိန်သတ်မှတ်ထားသော အညွှန်းကိန်း (Weighted Index Number)

ဤအညွှန်းကိန်းအမျိုးအစားသည် ပစ္စည်းတစ်ခုချင်းစီ၏ အရေးပါမှုကို အလေးချိန် (weight) သတ်မှတ်ခြင်းဖြင့် ပိုမိုတိကျစေသည်။ အလေးချိန်အတွက် ရောင်းအား သို့မဟုတ် စားသုံးမှုပမာဏကို အသုံးပြုနိုင်သည်။

၁.၅.၁ Laspeyres ၏ ဈေးနှုန်းအညွှန်းကိန်း (Laspeyres Price Index)

ဤနည်းလမ်းကို ပြင်သစ်စီးပွားရေးပညာရှင် Laspeyres က ၁၈၇၁ ခုနှစ်တွင် အဆိုပြုခဲ့သည်။ အခြေခံနှစ်တွင် စားသုံးခဲ့သော ကုန်ပစ္စည်းပမာဏကို အလေးချိန်အဖြစ် အသုံးပြုသည်။

Laspeyres အညွှန်းကိန်း (Laspeyres Index) = $\frac{\sum p_n q_0}{\sum p_0 q_0} \times 100$

- p_n = လက်ရှိကာလ၏ ဈေးနှုန်း
- p_0 = အခြေခံကာလ၏ ဈေးနှုန်း
- q_0 = အခြေခံကာလတွင် ရောင်းချခဲ့သော ပမာဏ

အနက်အဓိပ္ပာယ်-

ဤအညွှန်းကိန်းသည် အခြေခံကာလတွင် သတ်မှတ်ထားသော ပမာဏကို လက်ရှိကာလဈေးနှုန်းဖြင့် ဝယ်ယူရန် ကုန်ကျမည့်စရိတ်နှင့် အခြေခံကာလဈေးနှုန်းဖြင့် ဝယ်ယူရန် ကုန်ကျမည့်စရိတ်ကို နှိုင်းယှဉ်သည်။

အားသာချက်-

အခြေခံကာလနှင့် လက်ရှိကာလအပြင် အခြားကာလများနှင့်လည်း နှိုင်းယှဉ်နိုင်သည်။

အားနည်းချက်-

အခြေခံကာလ၏ ပမာဏသည် အချိန်ကြာလာသည်နှင့်အမျှ ခေတ်နောက်ကျသွားနိုင်သည်။

၁.၅.၂ Paasche ၏ ဈေးနှုန်းအညွှန်းကိန်း (Paasche Price Index)

ဤနည်းလမ်းကို ဂျာမန်စာရင်းအင်းပညာရှင် Paasche က ၁၈၇၄ ခုနှစ်တွင် အဆိုပြုခဲ့သည်။ လက်ရှိနှစ်၏ ပမာဏကို အလေးချိန်အဖြစ် အသုံးပြုသည်။

Paasche အညွှန်းကိန်း (Paasche Index) = $\frac{\sum p_n q_n}{\sum p_0 q_n} \times 100$

- q_n = လက်ရှိကာလတွင် ရောင်းချခဲ့သော ပမာဏ

၁.၅.၃ Irving Fisher ၏ အညွှန်းကိန်း (Irving Fisher's Index Number)

Fisher ၏ အညွှန်းကိန်းသည် Laspeyres နှင့် Paasche အညွှန်းကိန်းတို့၏ ဂျီသြမေတြီပျမ်းမျှ (Geometric Mean) ဖြစ်သည်။

Fisher ၏ အညွှန်းကိန်း (Fisher's Index) = $\sqrt{(Laspeyres\ Index) \times (Paasche\ Index)}$

ဤအညွှန်းကိန်းသည် အခြေခံနှစ်နှင့် လက်ရှိနှစ် နှစ်ခုလုံး၏ ပမာဏကို အသုံးပြုထားသောကြောင့် အကောင်းဆုံးဟု သတ်မှတ်ကြပြီး **Fisher's Ideal Index Number** ဟုလည်း ခေါ်သည်။

၁.၆ ကောင်းမွန်သော အညွှန်းကိန်း၏ စံနှုန်းများ (Criteria of a Good Index Number)

အညွှန်းကိန်းတစ်ခု၏ အရည်အသွေးကို စစ်ဆေးရန် စမ်းသပ်မှုများစွာ ရှိသည်။

၁။ **ယူနစ်စမ်းသပ်မှု (Unit Test):** အညွှန်းကိန်းသည် ဈေးနှုန်းနှင့် ပမာဏကို တိုင်းတာသည့် ယူနစ်ပေါ်တွင် မူတည်ခြင်း မရှိရ။
၂။ **အချိန်ပြောင်းပြန်စမ်းသပ်မှု (Time Reversal Test):** အညွှန်းကိန်းတွက်ချက်ရာတွင် အခြေခံကာလနှင့် လက်ရှိကာလကို ပြောင်းပြန်လှန်လိုက်လျှင်လည်း အချိုးသည် တူညီနေရမည်။

၁.၇ စားသုံးသူဈေးနှုန်းအညွှန်းကိန်း (Consumer Price Index - CPI)

စားသုံးသူဈေးနှုန်းအညွှန်းကိန်း (သို့) နေထိုင်စရိတ်အညွှန်းကိန်း (Cost of Living Index) သည် သတ်မှတ်ထားသော စားသုံးသူအုပ်စုတစ်ခုအတွက် ကုန်ပစ္စည်းအုပ်စုတစ်ခု၏ ဈေးနှုန်းပြောင်းလဲမှုကို တိုင်းတာသည်။

၎င်းသည် စားသုံးသူ၏ ဝယ်ယူနိုင်စွမ်းအား (purchasing power) ပြောင်းလဲမှုကို တိုင်းတာရန် အသုံးပြုသည်။ ဥပမာ- အလုပ်သမားများ၏ လုပ်ခလစာကို ငွေကြေးဖောင်းပွမှုနှင့်အညီ ပြင်ဆင်ရန် CPI ကို အသုံးပြုသည်။

CPI တွက်ချက်နည်း-

၁။ ပစ္စည်းတစ်ခုချင်းစီအတွက် ဈေးနှုန်းဆွေမျိုး (Price Relative) တွက်ပါ။
၂။ ထိုဆွေမျိုးများကို အခြေခံနှစ်၏ စားသုံးမှုပမာဏဖြင့် အလေးချိန်သတ်မှတ်ပါ။
၃။ အလေးချိန်သတ်မှတ်ထားသော ပျမ်းမျှကို တွက်ချက်ပါ။

၁.၈ အညွှန်းကိန်းများကို ဖောက်ပြန်ခြင်း (Deflating the Index Numbers)

ငွေကြေးဖောင်းပွမှုကြောင့် ဖြစ်ပေါ်လာသော ဈေးနှုန်းပြောင်းလဲမှုကို ပြင်ဆင်ခြင်းကို **Deflating** ဟုခေါ်သည်။

ဥပမာ- အခြေခံနှစ်တွင် တစ်ပြည်လျှင် ၁၀၀၀ ကျပ်ရှိသော ပစ္စည်းတစ်ခုသည် လက်ရှိနှစ်တွင် ၂၀၀၀ ကျပ်ဖြစ်လာပါက၊ ယခင်က ၁၀၀၀ ကျပ်ဖြင့် ဝယ်ယူနိုင်သော ပမာဏကို ယခုအခါ ၁၀၀၀ ကျပ်ဖြင့် ထက်ဝက်သာ ဝယ်ယူနိုင်တော့သည်။

အစစ်အမှန်လုပ်ခလစာ (Real Wage) = $\frac{\text{ငွေကြေးလုပ်ခလစာ (Nominal Wage)}}{\text{ဈေးနှုန်းအညွှန်းကိန်း (Price Index)}} \times 100$

ဤနည်းဖြင့် ငွေကြေးတန်ဖိုးကို ပြင်ဆင်ခြင်းဖြင့် အမှန်တကယ် ဝယ်ယူနိုင်စွမ်းအားကို တွက်ချက်နိုင်သည်။

ဤအခန်း၏ အဓိကအချက်မှာ အညွှန်းကိန်းများသည် ဈေးနှုန်းနှင့် ပမာဏပြောင်းလဲမှုကို တိုင်းတာရန် အရေးပါသော ကိရိယာများဖြစ်ပြီး ၎င်းတို့ကို မှန်ကန်စွာ တည်ဆောက်အသုံးပြုရန် သင့်လျော်သော နည်းလမ်းများကို သိရှိရန် လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၂) - ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာတွင် အရေးအပါဆုံး ဖြန့်ဝေမှုတစ်ခုဖြစ်သည့် **ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution)** အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၂.၁ နိဒါန်း (Introduction)

ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution) သည် လေ့လာတွေ့ရှိရလေ့ရှိသော စဉ်ဆက်မပြတ်ကျပန်းကိန်းရှင် (continuous random variable) တစ်မျိုးဖြစ်သည်။ ၎င်း၏ ဖြစ်နိုင်ခြေသိပ်သည်းဆဂရပ် (probability density function) သည် ခေါင်းလောင်းပုံသဏ္ဌာန် (bell-shaped curve) ရှိသည်။

ပုံမှန်ဖြန့်ဝေမှုသည် အောက်ပါတို့အတွက် အရေးပါသည်-

- ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်း (Statistical Inference) တွင် အဓိကအခန်းကဏ္ဍမှ ပါဝင်သည်။
- စီးပွားရေးဖြစ်စဉ်များစွာ (ဥပမာ - စတော့ဈေးကွက်၏ လစဉ်အမြတ်အစွန်း၊ ကုမ္ပဏီတစ်ခု၏ အပတ်စဉ်ရောင်းအား) သည် ပုံမှန်ဖြန့်ဝေမှုနှင့် အနီးစပ်ဆုံး ကိုက်ညီသည်။
- အလုပ်အကိုင်စွမ်းရည်စစ်ဆေးမှု (employment aptitude test) ရမှတ်များကိုလည်း ပုံမှန်ဖြန့်ဝေမှုဖြင့် ခန့်မှန်းနိုင်သည်။

ပုံမှန်ဖြန့်ဝေမှုသည် ၎င်း၏ ပျမ်းမျှ (mean) μ ကို ဗဟိုပြု၍ အချိုးကျ (symmetric) သည်။ ၎င်း၏ ပြန့်ကျဲမှု (spread) ကို စံသွေဖည်မှု (standard deviation) σ ဖြင့် သတ်မှတ်သည်။

၂.၂ ပုံမှန်ဖြန့်ဝေမှု၏ ဖြစ်နိုင်ခြေသိပ်သည်းဆလုပ်ဆောင်ချက် (The p.d.f. of the Normal)

ပုံမှန်ဖြန့်ဝေမှု၏ ဖြစ်နိုင်ခြေသိပ်သည်းဆလုပ်ဆောင်ချက် (Probability Density Function - pdf) ကို ဂျာမန်သချိုင်္ဂပညာရှင် **Gauss** (၁၇၇၇-၁၈၅၅) က ပထမဆုံး ရှာဖွေတွေ့ရှိခဲ့သည်။ ထို့ကြောင့် ပုံမှန်ဖြန့်ဝေမှုကို **Gaussian Distribution** ဟုလည်း ခေါ်သည်။

စံပုံမှန်ကိန်းရှင် (Standard Normal Variate) ၏ pdf-

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

ဤတွင် Z သည် ပျမ်းမျှ ၀ နှင့် စံသွေဖည်မှု ၁ ရှိသော ပုံမှန်ဖြန့်ဝေမှုကို ကိုယ်စားပြုသည်။ ၎င်းကို **Standard Normal Distribution** ဟုခေါ် ပြီး $Z \sim N(0,1)$ ဟု သတ်မှတ်သည်။

အရေးကြီးသော ဂုဏ်သတ္တိများ-

- $f(z)$ သည် သုညမဖြစ်နိုင်သော တန်ဖိုးမရှိပါ (ဆိုလိုသည်မှာ အလွန်ကြီးသော သို့မဟုတ် အလွန်သေးငယ်သော တန်ဖိုးများ ဖြစ်ပေါ် နိုင်သည်)။
- ဖြစ်နိုင်ခြေများကို တွက်ချက်ရန် စံပုံမှန်ဇယား (Standard Normal Table) ကို အသုံးပြုသည်။
- $\Phi(z)$ သည် စုစည်းဖြန့်ဝေမှုလုပ်ဆောင်ချက် (Cumulative Distribution Function - cdf) ကို ကိုယ်စားပြုသည်။

ဥပမာ-

$\Phi(1.0) = 0.8413$ ဆိုသည်မှာ ပျမ်းမျှထက် စံသွေဖည်မှု ၁ ခုအထက်တွင် ဒေတာ၏ ၈၄.၁၃% ရှိသည်ဟု ဆိုလိုသည်။

၂.၃ ပုံမှန် pdf များကို ပြောင်းလဲခြင်း (Transformation of Normal P.D.Fs)

မည်သည့် ပုံမှန်ကိန်းရှင်ကိုမဆို စံပုံမှန်ကိန်းရှင်အဖြစ် ပြောင်းလဲရန် အောက်ပါဖော်မြူလာကို အသုံးပြုသည်-

$$Z = \frac{X - \mu}{\sigma}$$

ဥပမာ-

ကြက်ဥတစ်လုံး၏ အလျားသည် ပျမ်းမျှ ၆ စင်တီမီတာနှင့် စံသွေဖည်မှု ၁.၄ စင်တီမီတာရှိသော ပုံမှန်ဖြန့်ဝေမှုတစ်ခုဖြစ်သည် ဆိုပါစို့။

(က) ကြက်ဥတစ်လုံး၏ အလျားသည် ၈ စင်တီမီတာထက် ပိုကြီးနိုင်ခြေကို ရှာရန်-

$$Z = \frac{8 - 6}{1.4} = 1.43$$

$\Phi(1.43) = 0.9236$ ဖြစ်သောကြောင့်၊

$$P(X > 8) = 1 - 0.9236 = 0.0764$$

(ခ) ကြက်ဥတစ်လုံး၏ အလျားသည် ၅ စင်တီမီတာထက် သေးငယ်နိုင်ခြေကို ရှာရန်-

$$Z = \frac{5 - 6}{1.4} = -0.714$$

$\Phi(0.714) = 0.7625$ ဖြစ်သောကြောင့်၊

$$P(X < 5) = 1 - 0.7625 = 0.2375$$

အထွေထွေဖော်မြူလာ-

$X \sim N(\mu, \sigma^2)$ ဖြစ်ပါက၊ $Z = \frac{X-\mu}{\sigma}$ သည် $Z \sim N(0, 1)$ ဖြစ်သည်။

၂.၄ အမှီအခိုကင်းသော ပုံမှန်ကျပန်းကိန်းရှင်များ၏ ပေါင်းလဒ် (Sums of Independent Normal Random Variables)

အကယ်၍ $X_1, X_2, \ldots, X_n$ တို့သည် အမှီအခိုကင်းသော ပုံမှန်ကျပန်းကိန်းရှင်များဖြစ်ပြီး ၎င်းတို့တွင် ပျမ်းမျှ $\mu_1, \mu_2, \ldots, \mu_n$ နှင့် ကွဲလွဲမှု $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ ရှိပါက၊ ၎င်းတို့၏ ပေါင်းလဒ် $Y = \sum_{i=1}^{n} c_i X_i$ သည် အောက်ပါအတိုင်း ပုံမှန်ဖြန့်ဝေမှုတစ်ခုဖြစ်သည်-

$$Y \sim N\left(\sum_{i=1}^{n} c_i, \mu_i \sum_{i=1}^{n} c_i^2 \sigma_i^2\right)$$

ဥပမာ ၂.၄.၁-

X_1 သည် ပျမ်းမျှ ၂ နှင့် ကွဲလွဲမှု ၃ ရှိသော ပုံမှန်ကိန်းရှင်ဖြစ်ပြီး X_2 သည် ပျမ်းမျှ ၁ နှင့် ကွဲလွဲမှု ၄ ရှိသော ပုံမှန်ကိန်းရှင်ဖြစ်သည်။ $Y = 2X_1 + 3X_2$ ၏ ဖြန့်ဝေမှုကို ရှာရန်-

$$Y \sim N\big(2, (2) + 3(1)2^2(3) + 3^2(4)\big) = N(7, 48)$$

အသုံးချမှု-

ဤသဘောတရားကို SAT စာမေးပွဲကဲ့သို့သော စံသတ်မှတ်စစ်ဆေးမှုများတွင် ကျောင်းသားတစ်ဦး၏ သန်ျာအမှတ်သည် အခြားကျောင်းသားတစ်ဦး၏ စကားပြေအမှတ်ထက် ပိုများနိုင်ခြေကို တွက်ချက်ရန် အသုံးပြုနိုင်သည်။

အချုပ်အားဖြင့်-

ပုံမှန်ဖြန့်ဝေမှုသည် ကိန်းဂဏာန်းပညာ၏ အခြေခံအုတ်မြစ်ဖြစ်ပြီး ဒေတာအများစုကို ခန့်မှန်းရန် အသုံးပြုသည်။ ၎င်းကို စံပုံမှန်ဖြန့်ဝေမှုသို့ ပြောင်းလဲခြင်းဖြင့် ဖြစ်နိုင်ခြေများကို အလွယ်တကူ တွက်ချက်နိုင်သည်။ ထို့အပြင် အမှီအခိုကင်းသော ပုံမှန်ကိန်းရှင်များ၏ ပေါင်းလဒ်သည်လည်း ပုံမှန်ဖြန့်ဝေမှုကို လိုက်နာသည်။

အခန်း (၃) - နမူနာကောက်ယူခြင်းနှင့် ခန့်မှန်းခြေ (Sampling and Estimation)

ဤအခန်းတွင် နမူနာကောက်ယူခြင်း (Sampling) နှင့် ခန့်မှန်းခြေ (Estimation) ဆိုင်ရာ အခြေခံသဘောတရားများ၊ နည်းလမ်းများနှင့် ယုံကြည်မှုကြားကာလ (Confidence Interval) တို့ကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၃.၁ နမူနာကောက်ယူခြင်း (Sampling)

နမူနာကောက်ယူခြင်း၏ ရည်ရွယ်ချက်မှာ လူဦးရေ (population) တစ်ခုလုံးကို မစစ်ဆေးဘဲ ၎င်း၏ အစိတ်အပိုင်း (နမူနာ) တစ်ခုကိုသာ စစ်ဆေးခြင်းဖြင့် လူဦးရေတစ်ခုလုံးအကြောင်း ကောက်ချက်ချရန်ဖြစ်သည်။

နမူနာကောက်ယူခြင်းသည် အောက်ပါအချက်များကြောင့် ဦးစားပေးသည်-

- ကုန်ကျစရိတ် (Cost)
- အချိန် (Time)
- ပျက်စီးစေသော စစ်ဆေးမှု (Destructive Testing)
- အကန့်အသတ်မရှိသော လူဦးရေ (Infinite Populations)

အဓိကအသုံးအနှုန်းများ-

- **ပါရာမီတာ (Parameter):** လူဦးရေတစ်ခု၏ ဂဏန်းဆိုင်ရာ တန်ဖိုး သို့မဟုတ် သွင်ပြင်လက္ခဏာ (ဥပမာ - လူဦးရေပျမ်းမျှ၊ အချိုးအစား)။
- **စာရင်းအင်းကိန်း (Statistic):** နမူနာတစ်ခု၏ ဂဏန်းဆိုင်ရာ အကျဉ်းချုပ် သို့မဟုတ် သွင်ပြင်လက္ခဏာ။

၃.၂ အမှတ်ခန့်မှန်းခြေ (Point Estimation)

ခန့်မှန်းခြေ (Estimation) သည် ကောက်ချက်ချသော ကိန်းဂဏန်းပညာ (Inferential Statistics) ၏ ပထမအပိုင်းဖြစ်သည်။ ခန့်မှန်းခြေ အမျိုးအစား နှစ်မျိုးရှိသည်-

၁။ **အမှတ်ခန့်မှန်းခြေ (Point Estimation)**

၂။ **ကြားကာလခန့်မှန်းခြေ (Interval Estimation)**

အမှတ်ခန့်မှန်းချက် (Point Estimate) ဆိုသည်မှာ လူဦးရေ၏ ပါရာမီတာကို ခန့်မှန်းရန် နမူနာမှ တွက်ချက်ထားသော စာရင်းအင်းကိန်း၏ တန်ဖိုးဖြစ်သည်။

၃.၂.၁ ပါရာမီတာများအတွက် ခန့်မှန်းချက်များအဖြစ် စာရင်းအင်းကိန်းများ (Statistics as Estimators for Parameters)

က) နမူနာအချိုးအစား (Sample Proportion)

သတ်မှတ်ထားသော သွင်ပြင်လက္ခဏာတစ်ခုရှိသော လူဦးရေတွင် အောင်မြင်မှု (success) နှင့် ကျရှုံးမှု (failure) ဟူ၍ ရလဒ်နှစ်မျိုးသာရှိသော စမ်းသပ်မှုကို **Bernoulli Experiment** ဟုခေါ်သည်။

နမူနာအချိုးအစား (Sample Proportion) ကို $\hat{p}$ ဖြင့် သတ်မှတ်သည်။

$$\hat{p} = \frac{x}{n}$$

- x = နမူနာတွင် သတ်မှတ်ထားသော သွင်ပြင်လက္ခဏာရှိသူဦးရေ
- n = နမူနာအရွယ်အစား

ဥပမာ ၃.၅-

ကျောင်းသား ၃၀ တွင် ဂဏန်းတွက်စက်ရှိသူ ၁၂ ဦးရှိပါက၊ နမူနာအချိုးအစားမှာ-

$$\hat{p} = \frac{12}{30} = 0.40 = 40\%$$

ခ) နမူနာပျမ်းမျှနှင့် နမူနာကွဲလွဲမှု (Sample Mean and Sample Variance)

လူဦးရေ၏ ပျမ်းမျှ μ နှင့် စံသွေဖည်မှု σ ကို ခန့်မှန်းရန် နမူနာပျမ်းမျှ $\bar{X}$ နှင့် နမူနာကွဲလွဲမှု S^2 တို့ကို အသုံးပြုသည်။

နမူနာပျမ်းမျှ (Sample Mean) -

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

နမူနာကွဲလွဲမှု (Sample Variance) -

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

ဥပမာ ၃.၆-

ကြွက် ၉ ကောင်၏ သွေးရည်ကြည်တွင် နိုက်ထရိုဂျင်ပါဝင်မှု (ဂရမ်/၁၀၀ စီစီ) သည်-

0.98, 0.83, 0.99, 0.86, 0.90, 0.81, 0.94, 0.92, 0.87

နမူနာပျမ်းမျှ-

$$\bar{X} = \frac{8.10}{9} = 0.90$$

နမူနာကွဲလွဲမှု-

$$S^2 = \frac{1}{8}\left(7.3220 - \frac{(8.10)^2}{9}\right) = 0.004$$
$$S = \sqrt{0.004} = 0.063$$

၃.၃ ကြားကာလခန့်မှန်းခြေ (Interval Estimation)

ကြားကာလခန့်မှန်းချက် (Interval Estimation) ဆိုသည်မှာ နမူနာမှ အချက်အလက်များကို အခြေခံ၍ လူဦးရေ၏ ပါရာမီတာ ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားတစ်ခုကို သတ်မှတ်ယုံကြည်မှုအဆင့် (confidence level) ဖြင့် တွက်ချက်ခြင်းဖြစ်သည်။

ကြားကာလခန့်မှန်းချက်ကို အောက်ပါပုံစံဖြင့် ဖော်ပြလေ့ရှိသည်-

$$\bar{X} \pm E$$

- E = အမှားအနားသတ် (Margin of Error)

၃.၄ ပါရာမီတာတစ်ခုအတွက် ယုံကြည်မှုကြားကာလ (Confidence Interval about One Parameter)

၃.၄.၁ အချိုးအစားတစ်ခုအတွက် ယုံကြည်မှုကြားကာလ (Confidence Interval about One Proportion)

လူဦးရေ၏ စစ်မှန်သော အချိုးအစား p အတွက် $100(1-\alpha)\%$ ယုံကြည်မှုကြားကာလမှာ-

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

ဥပမာ ၃.၇-

ဝန်ထမ်း ၅၀၀ တွင် နံနက်ခင်းသင်တန်းကို နှစ်သက်သူ ၁၆၀ ဦးရှိပါက ၉၅% ယုံကြည်မှုကြားကာလမှာ-

$$\hat{p} = \frac{160}{500} = 0.32,\ Z_{0.025} = 1.96$$

$$0.32 \pm 1.96\sqrt{\frac{0.32(0.68)}{500}} = 0.32 \pm 0.041$$

ဆိုလိုသည်မှာ စစ်မှန်သော အချိုးအစားသည် ၂၈% မှ ၃၆% ကြားတွင် ရှိသည်။

၃.၄.၂ ပျမ်းမျှတစ်ခုအတွက် ယုံကြည်မှုကြားကာလ (Confidence Interval about One Mean)

လူဦးရေ၏ ပျမ်းမျှ μ အတွက် ယုံကြည်မှုကြားကာလကို အောက်ပါအခြေအနေများဖြင့် တွက်ချက်သည်-

အခြေအနေ ၁- လူဦးရေ၏ ကွဲလွဲမှု (σ^2) ကို သိသော အခါ-

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

အခြေအနေ ၂- လူဦးရေ၏ ကွဲလွဲမှု (σ^2) ကို မသိသော်လည်း နမူနာအရွယ်အစား ကြီးသော အခါ ($n \geq 30$)-

$$\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

အခြေအနေ ၃- လူဦးရေ၏ ကွဲလွဲမှု (σ^2) ကို မသိဘဲ နမူနာအရွယ်အစား သေးငယ်သော အခါ (n < 30)-

ဤအခြေအနေတွင် t-ဖြန့်ဝေမှု (t-distribution) ကို အသုံးပြုသည်။

$$\bar{X} \pm t_{\alpha/2,(n-1)} \frac{S}{\sqrt{n}}$$

၃.၅ နမူနာအရွယ်အစား သတ်မှတ်ခြင်း (Sample Size Determination)

နမူနာအရွယ်အစားကို အောက်ပါဖော်မြူလာများဖြင့် တွက်ချက်နိုင်သည်-

ပျမ်းမျှအတွက်-

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E}\right)^2$$

အချိုးအစားအတွက်-

$$n = \frac{Z_{\alpha/2}^2 p(1-p)}{E^2}$$

- E = အမှားအနားသတ် (Margin of Error)

၃.၆ ယုံကြည်မှုကြားကာလနှင့် သည်းခံနိုင်မှုကြားကာလများ (Confidence Intervals vs Tolerance Intervals)

ယုံကြည်မှုကြားကာလ (Confidence Interval) သည် လူဦးရေ၏ ပါရာမီတာ (ဥပမာ - ပျမ်းမျှ) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို ခန့်မှန်းသည်။

သည်းခံနိုင်မှုကြားကာလ (Tolerance Interval) သည် လူဦးရေ၏ သတ်မှတ်ရာခိုင်နှုန်းတစ်ခု (ဥပမာ - ၉၉%) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို ခန့်မှန်းသည်။

အဓိကကွာခြားချက်-

- ယုံကြည်မှုကြားကာလသည် ပျမ်းမျှကိုသာ ခန့်မှန်းပြီး နမူနာအရွယ်အစား ကြီးလာသည်နှင့်အမျှ ၎င်း၏ အနံသည် သုည သို့ နီးကပ်လာသည်။
- သည်းခံနိုင်မှုကြားကာလသည် ဒေတာ၏ ပြန့်ကျဲမှုကို ထင်ဟပ်ပြီး နမူနာအရွယ်အစား ကြီးလာသည်နှင့်အမျှ လူဦးရေ၏ အမှန်တကယ်ပြန့်ကျဲမှုသို့ နီးကပ်လာသည်။

ဤအခန်း၏ အဓိကအချက်မှာ နမူနာကောက်ယူခြင်းနှင့် ခန့်မှန်းခြေသည် လူဦးရေတစ်ခုလုံးအကြောင်း ကောက်ချက်ချရန်အတွက် အရေးပါသော ကိရိယာများဖြစ်ပြီး ယုံကြည်မှုကြားကာလကို အသုံးပြုခြင်းဖြင့် ခန့်မှန်းချက်၏ တိကျမှုကို တိုင်းတာနိုင်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၄) - ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာ၏ ဒုတိယအဓိကအပိုင်းဖြစ်သည့် **ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing)** အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၄.၁ နိဒါန်း (Introduction)

ယူဆချက်စစ်ဆေးခြင်း (Statistical Hypothesis Testing) သည် ကောက်ချက်ချသော ကိန်းဂဏန်းပညာ (Inferential Statistics) ၏ ဒုတိယအဓိကအပိုင်းဖြစ်သည်။

စာရင်းအင်းဆိုင်ရာ ယူဆချက် (Statistical Hypothesis) ဆိုသည်မှာ လူဦးရေတစ်ခု သို့မဟုတ် တစ်ခုထက်ပိုသော လူဦးရေများ၏ ပါရာမီတာ (parameter) တစ်ခု သို့မဟုတ် တစ်ခုထက်ပိုသော ပါရာမီတာများအကြောင်း ပြုလုပ်ထားသော ယူဆချက် သို့မဟုတ် ထုတ်ပြန်ချက်တစ်ခုဖြစ်သည်။

ယူဆချက်တစ်ခုသည် မှန်နိုင်သည် သို့မဟုတ် မှားနိုင်သည်။ နမူနာမှ အချက်အလက်များကို အခြေခံ၍ ယူဆချက်မှန်သလား မှားသလားကို ဆုံးဖြတ်ရသည်။

၄.၂ အခြေခံသဘောတရားများ (Fundamental Concepts)

၄.၂.၁ သုညယူဆချက်နှင့် အစားထိုးယူဆချက် (Null and Alternative Hypotheses)

ယူဆချက်စစ်ဆေးခြင်း၏ ပထမအဆင့်မှာ အောက်ပါယူဆချက်နှစ်ခုကို သတ်မှတ်ခြင်းဖြစ်သည်-

၁။ **သုညယူဆချက် (Null Hypothesis - H_0):** ပြောင်းလဲမှုမရှိ၊ အကျိုးသက်ရောက်မှုမရှိ သို့မဟုတ် ကွာခြားမှုမရှိဟု ယူဆသော ယူဆချက်။
၂။ **အစားထိုးယူဆချက် (Alternative Hypothesis - H_1 သို့ H_a):** ပြောင်းလဲမှုရှိသည်ဟု ယူဆသော ယူဆချက်။

ဥပမာ-
ဆေးအသစ်တစ်မျိုးသည် ဆေးဟောင်းထက် ပိုမိုထိရောက်သည်ဟု ယူဆပါက-

- H_0: ဆေးအသစ်သည် ဆေးဟောင်းကဲ့သို့ ထိရောက်သည်။
- H_1: ဆေးအသစ်သည် ဆေးဟောင်းထက် ပိုမိုထိရောက်သည်။

၄.၂.၂ ဖြစ်နိုင်သော ဆုံးဖြတ်ချက်များ (Possible Decisions)

စစ်ဆေးမှုရလဒ်အရ အောက်ပါဆုံးဖြတ်ချက်များချနိုင်သည်-

၁။ **သုညယူဆချက် H_0 ကို ပယ်ချခြင်း (Reject H_0):** H_0 သည် မှားယွင်းသည်ဟု ကောက်ချက်ချပြီး H_1 ကို လက်ခံခြင်း။

၂။ **သုညယူဆချက် H_0 ကို ပယ်ချခြင်းမပြုခြင်း (Do Not Reject H_0):** နမူနာမှ H_0 ကို ပယ်ချရန် လုံလောက်သော အထောက်အထားမရှိခြင်း။

သတိပြုရန်- H_0 ကို ပယ်ချခြင်းမပြုခြင်းသည် H_0 မှန်သည်ဟု ဆိုလိုခြင်းမဟုတ်ပါ။ ၎င်းသည် နမူနာမှ လုံလောက်သော အထောက်အထား မတွေ့ရသေးခြင်းသာဖြစ်သည်။

၄.၂.၃ အမှားအမျိုးအစားများ (Types of Errors)

ယူဆချက်စစ်ဆေးခြင်းတွင် အမှားအမျိုးအစား နှစ်မျိုး ကျူးလွန်နိုင်သည်-

၁။ **ပထမအမျိုးအစားအမှား (Type I Error):** H_0 မှန်နေသော်လည်း ၎င်းကို ပယ်ချခြင်း။

- ဖြစ်နိုင်ခြေကို α (အယ်လ်ဖာ) ဖြင့် သတ်မှတ်သည်။
- $\alpha = P(\text{Type I Error}) = P(\text{Reject } H_0 \text{ when } H_0 \text{ is true})$

၂။ **ဒုတိယအမျိုးအစားအမှား (Type II Error):** H_0 မှားနေသော်လည်း ၎င်းကို ပယ်ချခြင်းမပြုခြင်း။

- ဖြစ်နိုင်ခြေကို β (ဘီတာ) ဖြင့် သတ်မှတ်သည်။
- $\beta = P(\text{Type II Error}) = P(\text{Do Not Reject } H_0 \text{ when } H_0 \text{ is false})$

ဇယား ၁- ဆုံးဖြတ်ချက်နှင့် အမှားအမျိုးအစားများ

ဆုံးဖြတ်ချက်	H_0 မှန်သည်	H_0 မှားသည်
H_0 ကို ပယ်ချသည်	**Type I Error** (α)	**Correct Decision**
H_0 ကို ပယ်ချခြင်းမပြု	**Correct Decision**	**Type II Error** (β)

စစ်ဆေးမှု၏ စွမ်းအား (Power of the Test): $1 - \beta$ သည် H_0 မှားနေချိန်တွင် ၎င်းကို ပယ်ချနိုင်သည့် ဖြစ်နိုင်ခြေဖြစ်သည်။

၄.၂.၄ အရေးပါမှုအဆင့် (Level of Significance)

α ကို **စစ်ဆေးမှု၏ အရေးပါမှုအဆင့် (Level of Significance)** ဟုခေါ်သည်။ အသုံးများသော တန်ဖိုးများမှာ 0.01, 0.05, နှင့် 0.10 ဖြစ်သည်။

၄.၂.၅ စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic)

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic) သည် နမူနာမှ ရရှိသော အချက်အလက်များပေါ်တွင် မူတည်ပြီး H_0 အောက်ရှိ ပါရာမီတာ(များ)၏ တန်ဖိုးနှင့် နှိုင်းယှဉ်သည်။

၄.၂.၆ ပယ်ချသည့်နယ်ပယ် (Critical Region or Rejection Region)

ပယ်ချသည့်နယ်ပယ် (Critical Region) သည် H_0 ကို ပယ်ချမည့် စာရင်းအင်းကိန်း၏ တန်ဖိုးများပါဝင်သော နယ်ပယ်ဖြစ်သည်။

၄.၂.၇ နိဂုံးချုပ်နှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုခြင်း (Conclusion and Interpretation)

- စာရင်းအင်းကိန်း၏ တန်ဖိုးသည် ပယ်ချသည့်နယ်ပယ်တွင် ကျရောက်ပါက H_0 ကို ပယ်ချသည်။
- မကျရောက်ပါက H_0 ကို ပယ်ချခြင်းမပြုပါ။

၄.၃ ပါရာမီတာတစ်ခုအကြောင်း ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing about One Parameter)

၄.၃.၁ အချိုးအစားတစ်ခုအကြောင်း ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing about One Proportion)

အဆင့်များ-

၁။ ယူဆချက်များ သတ်မှတ်ခြင်း-

- $H_0: p = p_0$ (Two-tailed)
- $H_0: p \geq p_0$ (Left-tailed)
- $H_0: p \leq p_0$ (Right-tailed)

၂။ အရေးပါမှုအဆင့် α သတ်မှတ်ခြင်းနှင့် ပယ်ချသည့်နယ်ပယ် သတ်မှတ်ခြင်း။

၃။ စစ်ဆေးမှုစာရင်းအင်းကိန်း တွက်ချက်ခြင်း-

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

၄။ ဆုံးဖြတ်ချက်ချခြင်းနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုခြင်း။

ဥပမာ ၄.၁-

နိုင်ငံခြားရင်းနှီးမြှုပ်နှံမှုရှိသော နိုင်ငံသားများ၏ ရာခိုင်နှုန်းသည် ၃၅% ထက် ပိုများသည်ဟု ယူဆပါက-

- $H_0: p \leq 0.35$
- $H_1: p > 0.35$
- $\alpha = 0.10$
- $n = 800, x = 320, \hat{p} = 0.40$
- $Z = \frac{0.40-0.35}{\sqrt{\frac{0.35(0.65)}{800}}} = 2.965$

$Z_{0.10} = 1.282$ ဖြစ်သောကြောင့် $Z = 2.965 > 1.282$ ဖြစ်သဖြင့် H_0 ကို ပယ်ချသည်။

၄.၃.၂ ပျမ်းမျှတစ်ခုအကြောင်း ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing about One Mean)

အခြေအနေ ၁- လူဦးရေ၏ကွဲလွဲမှု (σ^2) ကို သိသော အခါ-

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

အခြေအနေ ၂- လူဦးရေ၏ ကွဲလွဲမှု (σ^2) ကို မသိသော်လည်း နမူနာအရွယ်အစား ကြီးသော အခါ ($n \geq 30$)-

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

အခြေအနေ ၃- လူဦးရေ၏ ကွဲလွဲမှု (σ^2) ကို မသိဘဲ နမူနာအရွယ်အစား သေးငယ်သော အခါ (n < 30)-

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

t-ဖြန့်ဝေမှုကို $n - 1$ ဒီဂရီအက်ဆစ် (degrees of freedom) ဖြင့် အသုံးပြုသည်။

ဥပမာ ၄.၅-

ကြွက် ၂၅ ကောင်၏ ပျမ်းမျှအလေးချိန်သည် ၃၀ ဂရမ်ဟု ယူဆပါက-

- $H_0: \mu = 30$
- $H_1: \mu \neq 30$
- $\bar{X} = 33, S = 5, n = 25$
- $t = \frac{33-30}{5/\sqrt{25}} = 3$

$t_{0.025,24} = 2.064$ ဖြစ်သောကြောင့် $t = 3 > 2.064$ ဖြစ်သဖြင့် H_0 ကို ပယ်ချသည်။

၄.၃.၃ ကွဲလွဲမှုတစ်ခုအကြောင်း ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing about One Variance)

ကွဲလွဲမှုအတွက် Chi-Square (χ^2) စစ်ဆေးမှုကို အသုံးပြုသည်။

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

ဥပမာ ၄.၇-

$n = 15, S^2 = 3$ ဖြင့် $\sigma^2 > 1$ ဟု ယူဆပါက-

- $H_0: \sigma^2 \leq 1$
- $H_1: \sigma^2 > 1$
- $\chi^2 = \frac{(14)(3)}{1} = 42$
- $\chi^2_{0.05,14} = 23.685$

$\chi^2 = 42 > 23.685$ ဖြစ်သောကြောင့် H_0 ကို ပယ်ချသည်။

၄.၄ ပါရာမီတာနှစ်ခုအကြောင်း ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing about Two Parameters)

၄.၄.၁ အချိုးအစားနှစ်ခုအကြောင်း စစ်ဆေးခြင်း (Tests about Two Proportions)

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ (ပေါင်းစပ်ခန့်မှန်းချက်)

၄.၄.၂ ပျမ်းမျှနှစ်ခုအကြောင်း စစ်ဆေးခြင်း (Tests about Two Means)

ကွဲလွဲမှုများကို သိသော အခါ-

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

ကွဲလွဲမှုများကို မသိဘဲ နမူနာအရွယ်အစား ကြီးသော အခါ-

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

ကွဲလွဲမှုများ တူညီသည်ဟု ယူဆနိုင်သော အခါ-

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$ (ပေါင်းစပ်ကွဲလွဲမှု)

၄.၄.၃ ကွဲလွဲမှုနှစ်ခုအကြောင်း စစ်ဆေးခြင်း (Tests about Two Variances)

F-စစ်ဆေးမှုကို အသုံးပြုသည်။

$$F = \frac{S_1^2}{S_2^2}$$

၄.၅ Chi-Square စစ်ဆေးမှု (Chi-Square Test)

၄.၅.၁ အံဝင်ခွင်ကျဖြစ်မှုစစ်ဆေးခြင်း (Goodness of Fit Test)

လေ့လာထားသော အကြိမ်ရေ (Observed Frequency - O) နှင့် မျှော်လင့်ထားသော အကြိမ်ရေ (Expected Frequency - E) တို့ကို နှိုင်းယှဉ်သည်။

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

၄.၅.၂ အချက်အလက်ဇယားများ (Contingency Tables)

အမျိုးအစားနှစ်ခု သို့မဟုတ် နှစ်ခုထက်ပိုသော အမျိုးအစားများကြား အမှီအခိုကင်းမှုကို စစ်ဆေးရန် အသုံးပြုသည်။

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

- ဒီဂရီအက်ဆစ် = $(r-1)(c-1)$

ဤအခန်း၏ အဓိကအချက်မှာ ယူဆချက်စစ်ဆေးခြင်းသည် နမူနာမှ အချက်အလက်များကို အသုံးပြု၍ လူဦးရေတစ်ခုလုံးအကြောင်း ဆုံးဖြတ်ချက်ချရန်အတွက် အရေးပါသော ကိရိယာဖြစ်ပြီး မှန်ကန်သော ယူဆချက်များ၊ စစ်ဆေးမှုနည်းလမ်းများနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုချက်များ လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၅) - Regression Analysis

ဤအခန်းတွင် Regression Analysis ၏ အခြေခံသဘောတရား၊ ရိုးရှင်းသော ဆုတ်ယုတ်မှုပုံစံ (Simple Regression Model)၊ နှင့် ကိန်းရှင်များစွာပါဝင်သော ဆုတ်ယုတ်မှုပုံစံ (Multiple Regression Model) တို့ကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၅.၁ ရိုးရှင်းသော ဆုတ်ယုတ်မှုပုံစံ (A Simple Regression Model)

ဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်းသည် ကိန်းရှင်နှစ်ခု (သို့) နှစ်ခုထက်ပိုသော ကိန်းရှင်များကြား ဆက်နွယ်မှုကို လေ့လာရန် အသုံးပြုသည်။ ဥပမာ- အပူချိန်နှင့် ဓာတုပစ္စည်းထွက်ရှိမှုကြား ဆက်နွယ်မှုကို လေ့လာခြင်း။

ရိုးရှင်းသော ဆုတ်ယုတ်မှုပုံစံ (Simple Linear Regression Model):

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$

- Y_i = မှီခိုကိန်းရှင် (Dependent Variable) - ခန့်မှန်းလိုသော ကိန်းရှင်
- x_i = လွတ်လပ်ကိန်းရှင် (Independent Variable) - ကြိုတင်သိရှိထားသော ကိန်းရှင်
- α = Y-ကြားဖြတ် (Y-Intercept)
- β = ဆင်ခြေလျှော (Slope)
- ε_i = ကျပန်းအမှားကိန်း (Random Error Term) - $N(0, \sigma^2)$ ဖြန့်ဝေမှုရှိသည်။

အနည်းဆုံးနှစ်ထပ်ကိန်းနည်းလမ်း (Method of Least Squares):

အမှားကိန်းများ၏ နှစ်ထပ်ကိန်းပေါင်းလဒ် (Sum of Squared Errors - SSE) ကို အနည်းဆုံးဖြစ်စေသော α နှင့် β တန်ဖိုးများကို ရှာဖွေခြင်းဖြစ်သည်။

$$SSE = \sum_{i=1}^{n} (Y_i - \alpha - \beta x_i)^2$$

အနည်းဆုံးနှစ်ထပ်ကိန်း ခန့်မှန်းချက်များ (Least Squares Estimates):

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$$

ဥပမာ ၆.၅-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ကျောင်းသား ၁၀ ဦး၏ စာမေးပွဲအမှတ်များကို အခြေခံ၍ ဆုတ်ယုတ်မှုမျဉ်းကို တွက်ချက်ခြင်း။

၅.၂ ဆုတ်ယုတ်မှုပုံစံရှိ ပါရာမီတာများ၏ ဖြန့်ဝေမှု (Distributions of the Estimators)

$\hat{\alpha}$ ၏ ဖြန့်ဝေမှု-

- ပျမ်းမျှ = α
- ကွဲလွဲမှု = $\frac{\sigma^2}{n}$
- ပုံမှန်ဖြန့်ဝေမှု $N\left(\alpha, \frac{\sigma^2}{n}\right)$

$\hat{\beta}$ ၏ ဖြန့်ဝေမှု-

- ပျမ်းမျှ = β
- ကွဲလွဲမှု = $\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$
- ပုံမှန်ဖြန့်ဝေမှု $N\left(\beta, \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$

၅.၃ ယုံကြည်မှုကြားကာလနှင့် ယူဆချက်စစ်ဆေးခြင်း (Confidence Intervals and Hypothesis Testing)

β အတွက် ယုံကြည်မှုကြားကာလ-

$$\hat{\beta} \pm t_{\alpha/2,n-2} \cdot SE(\hat{\beta})$$

- $SE(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{\sum(x_i - \bar{x})^2}}$

- $\hat{\sigma}^2 = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}$

$\beta = 0$ အတွက် ယူဆချက်စစ်ဆေးခြင်း-

$$t = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

၅.၄ ခန့်မှန်းခြင်း (Prediction)

$x = x_0$ အတွက် Y ၏ ပျမ်းမျှတန်ဖိုး ခန့်မှန်းခြင်း-

$$\hat{Y}_0 = \hat{\alpha} + \hat{\beta}x_0$$

$x = x_0$ အတွက် ယုံကြည်မှုကြားကာလ-

$$\hat{Y}_0 \pm t_{\alpha/2,n-2} \cdot \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)}$$

$x = x_0$ အတွက် ခန့်မှန်းချက်ကြားကာလ (Prediction Interval)-

$$\hat{Y}_0 \pm t_{\alpha/2,n-2} \cdot \sqrt{\hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)}$$

၅.၅ ကိန်းရှင်များစွာပါဝင်သော ဆုတ်ယုတ်မှုပုံစံ (Multiple Regression Model)

ကိန်းရှင်များစွာပါဝင်သော ဆုတ်ယုတ်မှုပုံစံမှာ အောက်ပါအတိုင်းဖြစ်သည်-

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

ပုံမှန်ညီမျှခြင်းများ (Normal Equations):

ပါရာမီတာများကို ခန့်မှန်းရန် အောက်ပါညီမျှခြင်းများကို ဖြေရှင်းရသည်-

$$\sum_{i=1}^{n} Y_i = \hat{\beta}_0 n + \hat{\beta}_1 \sum X_{1i} + \cdots + \hat{\beta}_k \sum X_{ki}$$
$$\sum_{i=1}^{n} X_{1i} Y_i = \hat{\beta}_0 \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \cdots + \hat{\beta}_k \sum X_{1i} X_{ki}$$
$$\vdots$$
$$\sum_{i=1}^{n} X_{ki} Y_i = \hat{\beta}_0 \sum X_{ki} + \hat{\beta}_1 \sum X_{ki} X_{1i} + \cdots + \hat{\beta}_k \sum X_{ki}^2$$

ဥပမာ ၇.၆-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ဒေတာ ၅ ခုကို အခြေခံ၍ $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ ပုံစံကို ခန့်မှန်းခြင်း။

၅.၆ ဆုတ်ယုတ်မှုပုံစံရှိ ပါရာမီတာများအတွက် ယုံကြည်မှုကြားကာလများ (Confidence Intervals for Parameters in Regression Model)

β_j အတွက် ယုံကြည်မှုကြားကာလ-

$$\hat{\beta}_j \pm t_{\alpha/2,n-k-1} \cdot SE(\hat{\beta}_j)$$

ဤအခန်း၏ အဓိကအချက်မှာ ဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်းသည် ကိန်းရှင်များကြား ဆက်နွယ်မှုကို လေ့လာရန် အရေးပါသော ကိရိယာဖြစ်ပြီး ရိုးရှင်းသော ဆုတ်ယုတ်မှုပုံစံမှ စတင်ကာ ကိန်းရှင်များစွာပါဝင်သော ပုံစံအထိ တိုးချဲ့နိုင်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၆) - အမှတ်ခန့်မှန်းခြင်း (Point Estimation)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာ၏ အခြေခံအကျဆုံး
ခန့်မှန်းခြေနည်းလမ်းများဖြစ်သည့် **အမှတ်ခန့်မှန်းခြင်း (Point Estimation)** အကြောင်းကို
အသေးစိတ် ရှင်းလင်းထားပါသည်။

၆.၁ နိဒါန်း (Introduction)

ကိန်းဂဏန်းပညာတွင် ခန့်မှန်းခြေ (Estimation) သည် အရေးပါသော
အစိတ်အပိုင်းတစ်ခုဖြစ်သည်။ ကျွန်ုပ်တို့သည် လူဦးရေ (Population) တစ်ခု၏
သွင်ပြင်လက္ခဏာများ (Parameters) ကို နမူနာ (Sample) မှ ရရှိသော အချက်အလက်များဖြင့်
ခန့်မှန်းလေ့ရှိသည်။

ခန့်မှန်းခြေ အမျိုးအစား နှစ်မျိုး-

၁။ **အမှတ်ခန့်မှန်းခြေ (Point Estimation)** - ပါရာမီတာတစ်ခုကို တန်ဖိုးတစ်ခုတည်းဖြင့်
ခန့်မှန်းခြင်း။
၂။ **ကြားကာလခန့်မှန်းခြေ (Interval Estimation)** - ပါရာမီတာတစ်ခုကို
တန်ဖိုးအပိုင်းအခြားတစ်ခုဖြင့် ခန့်မှန်းခြင်း။

ဤအခန်းတွင် အမှတ်ခန့်မှန်းခြေကိုသာ အဓိကထား ရှင်းလင်းထားပါသည်။

၆.၂ ဆက်စပ်ကိန်းဂဏန်းများ (Descriptive Statistics)

ဒေတာများကို ဖော်ပြရန် အောက်ပါ ကိန်းဂဏန်းများကို အသုံးပြုသည်-

နမူနာပျမ်းမျှ (Sample Mean):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

နမူနာကွဲလွဲမှု (Sample Variance):

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

နမူနာစံသွေဖည်မှု (Sample Standard Deviation):

$$s = \sqrt{s^2}$$

ဥပမာ ၆.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
သကြားလုံး ၄၀ ၏ အလေးချိန်များကို အခြေခံ၍ နမူနာပျမ်းမျှနှင့် နမူနာကွဲလွဲမှုကို တွက်ချက်ခြင်း။

၆.၃ အမှတ်ခန့်မှန်းခြေ (Point Estimation)

အမှတ်ခန့်မှန်းခြေတွင် အောက်ပါအချက်များကို သိရှိရန် လိုအပ်သည်-

၁။ **ခန့်မှန်းချက်၏ မျှော်လင့်ချက် (Expected Value of the Estimator)**
၂။ **ခန့်မှန်းချက်၏ ကွဲလွဲမှု (Variance of the Estimator)**
၃။ **ခန့်မှန်းချက်၏ ဖြန့်ဝေမှု (Distribution of the Estimator)**

၆.၄ ခန့်မှန်းချက်၏ ဂုဏ်သတ္တိများ (Properties of Estimators)

မျှခြေမဲ့ခန့်မှန်းချက် (Unbiased Estimator):

ခန့်မှန်းချက်တစ်ခု၏ မျှော်လင့်ချက်သည် ခန့်မှန်းလိုသော ပါရာမီတာနှင့် ညီမျှပါက ၎င်းကို **မျှခြေမဲ့ခန့်မှန်းချက် (Unbiased Estimator)** ဟုခေါ်သည်။

$$E(\hat{\theta}) = \theta$$

ထိရောက်မှု (Efficiency):

ကွဲလွဲမှုအနည်းဆုံးရှိသော ခန့်မှန်းချက်ကို **ထိရောက်သော ခန့်မှန်းချက် (Efficient Estimator)** ဟုခေါ်သည်။

ကိုက်ညီမှု (Consistency):

နမူနာအရွယ်အစား ကြီးလာသည်နှင့်အမျှ ခန့်မှန်းချက်သည် ပါရာမီတာ၏ စစ်မှန်သောတန်ဖိုးသို့ နီးကပ်လာပါက ၎င်းကို **ကိုက်ညီသော ခန့်မှန်းချက် (Consistent Estimator)** ဟုခေါ်သည်။

၆.၅ ခန့်မှန်းခြေနည်းလမ်းများ (Methods of Estimation)

၆.၅.၁ များများဖြစ်နိုင်ဆုံး ခန့်မှန်းခြေနည်းလမ်း (Method of Maximum Likelihood - MLE)

ဤနည်းလမ်းတွင် နမူနာမှ လေ့လာတွေ့ရှိရသော ဒေတာများကို ဖြစ်ပေါ်စေနိုင်ခြေ အများဆုံးဖြစ်စေသော ပါရာမီတာ၏ တန်ဖိုးကို ရှာဖွေသည်။

လုပ်ဆောင်ချက်-

$$L(\theta) = \prod_{i=1}^{n} f\ (x_i; \theta)$$

MLE ရှာဖွေနည်း-

၁။ လုပ်ဆောင်ချက် $L(\theta)$ ကို ရေးပါ။

၂။ $\ln L\ (\theta)$ ကို ရေးပါ။

၃။ $\frac{d}{d\theta} \ln L\ (\theta) = 0$ ကို ဖြေရှင်းပါ။

ဥပမာ ၆.၄-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ထပ်ကိန်းဖြန့်ဝေမှု (Exponential Distribution) အတွက် MLE ရှာဖွေခြင်း။

ဥပမာ ၆.၄-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution) အတွက် MLE ရှာဖွေခြင်း။

၆.၅.၂ လဟုတ်စဉ်ခန့်မှန်းခြေနည်းလမ်း (Method of Moments - MOM)

ဤနည်းလမ်းတွင် လူဦးရေ၏ လဟုတ်စဉ်များ (Population Moments) ကို နမူနာ၏ လဟုတ်စဉ်များ (Sample Moments) နှင့် ညီမျှသည်ဟု ယူဆသည်။

ဥပမာ ၆.၄-၇ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution) အတွက် MOM ရှာဖွေခြင်း။

၆.၆ ခန့်မှန်းချက်၏ အမှားအနားသတ် (Standard Error of the Estimator)

ခန့်မှန်းချက်၏ စံသွေဖည်မှုကို **အမှားအနားသတ် (Standard Error)** ဟုခေါ်သည်။

$$SE(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$$

၆.၇ လုံလောက်သော စာရင်းအင်းကိန်းများ (Sufficient Statistics)

လုံလောက်သော စာရင်းအင်းကိန်း (Sufficient Statistic) ဆိုသည်မှာ နမူနာတွင်ပါဝင်သော ပါရာမီတာအကြောင်း အချက်အလက်အားလုံးကို ပိုင်ဆိုင်ထားသော စာရင်းအင်းကိန်းဖြစ်သည်။

Fisher-Neyman Factorization Theorem-
$T(X)$ သည် θ အတွက် လုံလောက်သော စာရင်းအင်းကိန်းဖြစ်ပါက၊

$$f(x_1, x_2, \dots, x_n, \theta) = g(T; (x)\theta) \cdot h(x_1, x_2, \dots, x_n)$$

၆.၈ ဘိုင်းစ်ရှန်း ခန့်မှန်းခြေ (Bayesian Estimation)

ဘိုင်းစ်ရှန်းခန့်မှန်းခြေတွင် ပါရာမီတာ θ ကို ကျပန်းကိန်းရှင်အဖြစ် သတ်မှတ်ပြီး ၎င်း၏ ကြိုတင်ဖြန့်ဝေမှု (Prior Distribution) $\pi(\theta)$ ကို သတ်မှတ်သည်။

နောက်ဆက်တွဲ ဖြန့်ဝေမှု (Posterior Distribution):

$$\pi(\theta \mid x) = \frac{f(x \mid \theta)\pi(\theta)}{\int f\,(x \mid \theta)\pi(\theta)d\theta}$$

ဘိုင်းစ်ရှန်း ခန့်မှန်းချက် (Bayesian Estimator):

နောက်ဆက်တွဲ ဖြန့်ဝေမှု၏ ပျမ်းမျှ (Mean of Posterior Distribution) ကို ခန့်မှန်းချက်အဖြစ် အသုံးပြုသည်။

$$\hat{\theta}_{Bayes} = E(\theta \mid x) = \int \theta \, \pi(\theta \mid x) d\theta$$

ဤအခန်း၏ အဓိကအချက်မှာ အမှတ်ခန့်မှန်းခြေသည် လူဦးရေ၏ ပါရာမီတာများကို နမူနာမှ ခန့်မှန်းရန် အခြေခံနည်းလမ်းဖြစ်ပြီး မျှခြေမဲ့ခန့်မှန်းချက်၊ ထိရောက်မှု၊ ကိုက်ညီမှုစသည့် ဂုဏ်သတ္တိများ၊ နှင့် ခန့်မှန်းခြေနည်းလမ်းအမျိုးမျိုးကို သိရှိရန် လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၇) - ကြားကာလခန့်မှန်းခြေ (Interval Estimation)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာ၏ အရေးပါသော အပိုင်းတစ်ခုဖြစ်သည့် **ကြားကာလခန့်မှန်းခြေ (Interval Estimation)** အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ ယုံကြည်မှုကြားကာလ (Confidence Interval) တည်ဆောက်ပုံ၊ အဓိပ္ပာယ်ဖွင့်ဆိုပုံနှင့် အသုံးချမှုများကို ဖော်ပြထားပါသည်။

၇.၁ ပျမ်းမျှအတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for Means)

ယုံကြည်မှုကြားကာလ (Confidence Interval) ဆိုသည်မှာ လူဦးရေ၏ ပါရာမီတာတစ်ခု (ဥပမာ - ပျမ်းမျှ) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို သတ်မှတ်ယုံကြည်မှုအဆင့် (Confidence Level) ဖြင့် ခန့်မှန်းခြင်းဖြစ်သည်။

၇.၁.၁ ကွဲလွဲမှု (σ^2) ကို သိသော အခါ (When Variance is Known)

လူဦးရေ၏ ကွဲလွဲမှု σ^2 ကို သိပါက၊ ပျမ်းမျှ μ အတွက် $100(1-\alpha)\%$ ယုံကြည်မှုကြားကာလမှာ-

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

ဥပမာ ၇.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
မီးသီး ၂၇ လုံး၏ ပျမ်းမျှသက်တမ်းအတွက် ၉၅% ယုံကြည်မှုကြားကာလ တွက်ချက်ခြင်း။

၇.၁.၂ ကွဲလွဲမှု (σ^2) ကို မသိသော အခါ (When Variance is Unknown)

လူဦးရေ၏ ကွဲလွဲမှု σ^2 ကို မသိပါက၊ နမူနာကွဲလွဲမှု S^2 ကို အသုံးပြုပြီး t-ဖြန့်ဝေမှုဖြင့် ယုံကြည်မှုကြားကာလကို တွက်ချက်သည်။

$$\bar{X} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

ဥပမာ ၇.၁-၅ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
နွားနို့ထွက်ရှိမှုအတွက် ၉၀% ယုံကြည်မှုကြားကာလ တွက်ချက်ခြင်း။

၇.၂ ပျမ်းမျှနှစ်ခုကြား ကွာခြားချက်အတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for the Difference of Two Means)

၇.၂.၁ ကွဲလွဲမှုများကို သိသော အခါ (When Variances are Known)

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

၇.၂.၂ ကွဲလွဲမှုများကို မသိဘဲ နမူနာအရွယ်အစား ကြီးသော အခါ (When Variances are Unknown but Sample Sizes are Large)

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

၇.၂.၃ ကွဲလွဲမှုများ တူညီသည်ဟု ယူဆနိုင်သော အခါ (When Variances are Assumed Equal)

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, n_1+n_2-2} \cdot S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ (ပေါင်းစပ်ကွဲလွဲမှု)

၇.၂.၄ ကွဲလွဲမှုများ မတူညီသော အခါ (When Variances are Not Equal)

Welch ၏ အနီးစပ်ဆုံး t-စစ်ဆေးမှုကို အသုံးပြုသည်။

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

ဒီဂရီအက်ဆစ် (Degrees of Freedom) ကို အောက်ပါအတိုင်း တွက်ချက်သည်-

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

၇.၂.၅ မှီခိုနေသော နမူနာများအတွက် (For Dependent Samples)

နမူနာနှစ်ခုသည် မှီခိုနေပါက (ဥပမာ - တူညီသော လူတစ်ဦးကို အကြိမ်နှစ်ကြိမ် တိုင်းတာခြင်း)၊ ကွာခြားချက် $D_i = X_{1i} - X_{2i}$ ကို အသုံးပြုသည်။

\bar{D} \pm t_{\alpha/2, n-1} \frac{S_D}{\sqrt{n}} ``` --- ## ၇.၃ အချိုးအစားအတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for Proportions) ### ၇.၃.၁ အချိုးအစားတစ်ခုအတွက် (For a Single Proportion) \[\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}

ဥပမာ ၇.၃-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
သကြားလုံးအလေးချိန်အတွက် ၉၀% ယုံကြည်မှုကြားကာလ တွက်ချက်ခြင်း။

၇.၃.၂ အချိုးအစားနှစ်ခုကြား ကွာခြားချက်အတွက် (For the Difference of Two Proportions)

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

၇.၄ နမူနာအရွယ်အစား သတ်မှတ်ခြင်း (Sample Size Determination)

၇.၄.၁ ပျမ်းမျှအတွက် (For a Mean)

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E} \right)^2$$

- E = အမှားအနားသတ် (Margin of Error)

၇.၄.၂ အချိုးအစားအတွက် (For a Proportion)

$$n = \frac{Z_{\alpha/2}^2 p(1-p)}{E^2}$$

အကယ်၍ p ၏ တန်ဖိုးကို မသိပါက $p = 0.5$ ကို အသုံးပြုနိုင်သည် (၎င်းသည် ကွဲလွဲမှုကို အမြင့်ဆုံးဖြစ်စေသောကြောင့် ရှေးရိုးဆန်သော နည်းလမ်းဖြစ်သည်)။

၇.၅ ရာခိုင်နှုန်းများအတွက် ဖြန့်ဝေမှုမပါသော ယုံကြည်မှုကြားကာလများ (Distribution-Free Confidence Intervals for Percentiles)

ဤနည်းလမ်းသည် လူဦးရေ၏ ဖြန့်ဝေမှုပုံစံကို မသိဘဲ ရာခိုင်နှုန်းများ (Percentiles) အတွက် ယုံကြည်မှုကြားကာလကို တည်ဆောက်ရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ နမူနာကို စီရန်လိုအပ်သည်။
၂။ သင့်လျော်သော အစီအစဉ်စာရင်းအင်းကိန်းများ (Order Statistics) ကို ရွေးချယ်သည်။
၃။ နှစ်ထပ်ကိန်းဖြန့်ဝေမှု (Binomial Distribution) ကို အသုံးပြု၍ ယုံကြည်မှုအဆင့်ကို တွက်ချက်သည်။

၇.၆ ယုံကြည်မှုကြားကာလများနှင့် သည်းခံနိုင်မှုကြားကာလများ (Confidence Intervals vs Tolerance Intervals)

ယုံကြည်မှုကြားကာလ (Confidence Interval):

- လူဦးရေ၏ ပါရာမီတာ (ဥပမာ - ပျမ်းမျှ) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို ခန့်မှန်းသည်။
- နမူနာအရွယ်အစား ကြီးလာသည်နှင့်အမျှ အနံသည် သုညသို့ နီးကပ်လာသည်။

သည်းခံနိုင်မှုကြားကာလ (Tolerance Interval):

- လူဦးရေ၏ သတ်မှတ်ရာခိုင်နှုန်းတစ်ခု (ဥပမာ - ၉၉%) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို ခန့်မှန်းသည်။
- နမူနာအရွယ်အစား ကြီးလာသည်နှင့်အမျှ လူဦးရေ၏ အမှန်တကယ်ပြန့်ကျဲမှုသို့ နီးကပ်လာသည်။

ဤအခန်း၏ အဓိကအချက်မှာ ကြားကာလခန့်မှန်းခြေသည် လူဦးရေ၏ ပါရာမီတာများကို တန်ဖိုးအပိုင်းအခြားဖြင့် ခန့်မှန်းရန် အရေးပါသောကိရိယာဖြစ်ပြီး ယုံကြည်မှုအဆင့်၊ နမူနာအရွယ်အစားနှင့် အမှားအနားသတ်တို့ကို ထည့်သွင်းစဉ်းစားရန် လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၈) - ယူဆချက်စစ်ဆေးခြင်း (Tests of Statistical Hypotheses)

ဤအခန်းတွင် ယူဆချက်စစ်ဆေးခြင်း၏ အခြေခံသဘောတရားများ၊ စစ်ဆေးမှုအဆင့်များ၊ နှင့် အမျိုးအစားအမျိုးမျိုးကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၈.၁ ပျမ်းမျှတစ်ခုအကြောင်း စစ်ဆေးခြင်း (Tests About One Mean)

ယူဆချက်စစ်ဆေးခြင်းတွင် အောက်ပါအဆင့်များ ပါဝင်သည်-

၁။ သုညယူဆချက် (Null Hypothesis) H_0 နှင့် အစားထိုးယူဆချက် (Alternative Hypothesis) H_1 ကို သတ်မှတ်ခြင်း။

၂။ အရေးပါမှုအဆင့် (Level of Significance) α ကို ရွေးချယ်ခြင်း။

၃။ သင့်လျော်သော စာရင်းအင်းစစ်ဆေးမှု (Test Statistic) ကို တွက်ချက်ခြင်း။

၄။ ပယ်ချသည့်နယ်ပယ် (Critical Region) ကို သတ်မှတ်ခြင်း။

၅။ ဆုံးဖြတ်ချက်ချခြင်းနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုခြင်း။

၈.၁.၁ ကွဲလွဲမှု (σ^2) ကို သိသော အခါ (When Variance is Known)

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

ဆုံးဖြတ်ချက်စည်းမျဉ်း (Decision Rule):

ယူဆချက်	ပယ်ချသည့်နယ်ပယ်		
$H_1: \mu > \mu_0$	$Z > Z_\alpha$		
$H_1: \mu < \mu_0$	$Z < -Z_\alpha$		
$H_1: \mu \neq \mu_0$	(	Z	> Z_{\alpha/2})

ဥပမာ ၈.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

သံမဏိဘားတစ်ခု၏ ဆန့်နိုင်အားကို စစ်ဆေးခြင်း။

၈.၁.၂ ကွဲလွဲမှု (σ^2) ကို မသိသော အခါ (When Variance is Unknown)

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

ဆုံးဖြတ်ချက်စည်းမျဉ်း (Decision Rule):

ယူဆချက်	ပယ်ချသည့်နယ်ပယ်		
$H_1: \mu > \mu_0$	$t > t_{\alpha,n-1}$		
$H_1: \mu < \mu_0$	$t < -t_{\alpha,n-1}$		
$H_1: \mu \neq \mu_0$	(	t	> t_{\alpha/2, n-1})

ဥပမာ ၈.၁-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ကြွက်တစ်ကောင်၏ အကျိတ်ကြီးထွားမှုကို စစ်ဆေးခြင်း။

၈.၂ ပျမ်းမျှနှစ်ခုကြား ကွာခြားချက်ကို စစ်ဆေးခြင်း (Tests of the Equality of Two Means)

၈.၂.၁ ကွဲလွဲမှုများကို သိသော အခါ (When Variances are Known)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

၈.၂.၂ ကွဲလွဲမှုများကို မသိဘဲ နမူနာအရွယ်အစား ကြီးသော အခါ (When Variances are Unknown but Sample Sizes are Large)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

၈.၂.၃ ကွဲလွဲမှုများ တူညီသည်ဟု ယူဆနိုင်သော အခါ (When Variances are Assumed Equal)

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

ဥပမာ ၈.၂-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ပဲပင်ကြီးထွားမှုကို စစ်ဆေးခြင်း။

၈.၃ အချိုးအစားများအကြောင်း စစ်ဆေးခြင်း (Tests About Proportions)

၈.၃.၁ အချိုးအစားတစ်ခုအတွက် (For a Single Proportion)

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

၈.၃.၂ အချိုးအစားနှစ်ခုကြား ကွာခြားချက်အတွက် (For the Difference of Two Proportions)

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- $\hat{p} = \frac{x_1+x_2}{n_1+n_2}$

ဥပမာ ၈.၃-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
အံစာတုံးများ၏ တန်းတူညီမျှမှုကို စစ်ဆေးခြင်း။

၈.၄ Wilcoxon စစ်ဆေးမှုများ (The Wilcoxon Tests)

Wilcoxon စစ်ဆေးမှုများသည် ဖြန့်ဝေမှုပုံစံကို မသိဘဲ အသုံးပြုနိုင်သော စစ်ဆေးမှုများ (Non-parametric Tests) ဖြစ်သည်။

၈.၄.၁ Wilcoxon Signed-Rank စစ်ဆေးမှု (Wilcoxon Signed-Rank Test)

ဤစစ်ဆေးမှုသည် ပျမ်းမျှ (median) တစ်ခုကို စစ်ဆေးရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ ကွာခြားချက်များ $D_i = X_i - \mu_0$ ကို တွက်ချက်ပါ။

၂။ $|D_i|$ ကို အစဉ်လိုက်စီပါ။

၃။ အဆင့်သတ်မှတ်ပါ (Rank)။

၄။ အပြုသဘောဆောင်သော အဆင့်များ၏ ပေါင်းလဒ် W ကို တွက်ချက်ပါ။

စ.၄.၂ Wilcoxon Rank-Sum စစ်ဆေးမှု (Wilcoxon Rank-Sum Test)

ဤစစ်ဆေးမှုသည် လွတ်လပ်သော နမူနာနှစ်ခု၏ ပျမ်းမျှများကို နှိုင်းယှဉ်ရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ နမူနာနှစ်ခုကို ပေါင်းစပ်ပြီး အစဉ်လိုက်စီပါ။

၂။ အဆင့်သတ်မှတ်ပါ (Rank)။

၃။ နမူနာတစ်ခု၏ အဆင့်ပေါင်းလဒ် W ကို တွက်ချက်ပါ။

စ.၅ စစ်ဆေးမှု၏ စွမ်းအား (Power of a Statistical Test)

စစ်ဆေးမှု၏ စွမ်းအား (Power of a Test) ဆိုသည်မှာ သုညယူဆချက် H_0 မှားနေချိန်တွင် ၎င်းကို ပယ်ချနိုင်သည့် ဖြစ်နိုင်ခြေဖြစ်သည်။

$$\text{Power} = 1 - \beta = P(\text{Reject } H_0 \mid H_0 \text{ is false})$$

စွမ်းအားလုပ်ဆောင်ချက် (Power Function):

$$K(\theta) = P(\text{Reject } H_0 \mid \theta)$$

စ.၆ အကောင်းဆုံးပယ်ချသည့်နယ်ပယ်များ (Best Critical Regions)

Neyman-Pearson Lemma-

$H_0: \theta = \theta_0$ နှင့် $H_1: \theta = \theta_1$ အတွက် အကောင်းဆုံးပယ်ချသည့်နယ်ပယ်မှာ-

$$\frac{L(\theta_0)}{L(\theta_1)} \leq k$$

- $L(\theta)$ = များများဖြစ်နိုင်ဆုံးလုပ်ဆောင်ချက် (Likelihood Function)
- k = သတ်မှတ်ထားသော ကိန်းသေ

၈.၇ များများဖြစ်နိုင်ဆုံးအချိုးစစ်ဆေးမှု (Likelihood Ratio Tests)

များများဖြစ်နိုင်ဆုံးအချိုး (Likelihood Ratio):

$$\lambda = \frac{\max_{\theta \in \omega} L(\theta)}{\max_{\theta \in \Theta} L(\theta)}$$

- ω = H_0 အောက်ရှိ ပါရာမီတာနယ်ပယ်
- Θ = စုစုပေါင်းပါရာမီတာနယ်ပယ်

λ သည် သေးငယ်ပါက H_0 ကို ပယ်ချသည်။

ဤအခန်း၏ အဓိကအချက်မှာ ယူဆချက်စစ်ဆေးခြင်းသည် နမူနာမှ အချက်အလက်များကို အသုံးပြု၍ လူဦးရေတစ်ခုလုံးအကြောင်း ဆုံးဖြတ်ချက်ချရန်အတွက် အရေးပါသော ကိရိယာဖြစ်ပြီး မှန်ကန်သော ယူဆချက်များ၊ သင့်လျော်သော စစ်ဆေးမှုများနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုချက်များ လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၉) - နောက်ထပ်စစ်ဆေးမှုများ (More Tests)

ဤအခန်းတွင် Chi-Square စစ်ဆေးမှုများ၊ ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (Analysis of Variance - ANOVA)၊ နှင့် စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်း (Statistical Quality Control) တို့ကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၉.၁ Chi-Square Goodness-of-Fit စစ်ဆေးမှု (Chi-Square Goodness-of-Fit Test)

ဤစစ်ဆေးမှုသည် လေ့လာထားသော အကြိမ်ရေ (Observed Frequency) နှင့် မျှော်လင့်ထားသော အကြိမ်ရေ (Expected Frequency) တို့ကို နှိုင်းယှဉ်ခြင်းဖြင့် ဒေတာတစ်ခုသည် သတ်မှတ်ထားသော ဖြန့်ဝေမှုတစ်ခုနှင့် ကိုက်ညီမှုရှိမရှိကို စစ်ဆေးသည်။

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

- O_i = အုပ်စု i တွင် လေ့လာရရှိသော အကြိမ်ရေ
- E_i = အုပ်စု i တွင် မျှော်လင့်ထားသော အကြိမ်ရေ
- k = အုပ်စုအရေအတွက်

ဒီဂရီအက်ဆစ် (Degrees of Freedom): $df = k - 1$ (ပါရာမီတာများကို ခန့်မှန်းထားပါက $df = k - 1 - m$)

ဥပမာ ၉.၁-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ဒင်္ဂါးပစ်စမ်းသပ်မှုတစ်ခု၏ ရလဒ်များသည် Binomial ဖြန့်ဝေမှုနှင့် ကိုက်ညီမှုရှိမရှိကို စစ်ဆေးခြင်း။

၉.၂ ဆက်စပ်ဇယားများ (Contingency Tables)

ဆက်စပ်ဇယားများသည် အမျိုးအစားနှစ်ခု သို့မဟုတ် နှစ်ခုထက်ပိုသော အမျိုးအစားများကြား အမှီအခိုကင်းမှုကို စစ်ဆေးရန် အသုံးပြုသည်။

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

- O_{ij} = အတန်း i နှင့် ကော်လံ j တွင် လေ့လာရရှိသော အကြိမ်ရေ
- $E_{ij} = \frac{(Row\ Total_i)(Column\ Total_j)}{Grand\ Total}$
- r = အတန်းအရေအတွက်
- c = ကော်လံအရေအတွက်

ဒီဂရီအက်ဆစ် (Degrees of Freedom): $df = (r-1)(c-1)$

ဥပမာ ၉.၂-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ကျောင်းသားများ၏ ကျားမနှင့် ဌာနရွေးချယ်မှုကြား အမှီအခိုကင်းမှုကို စစ်ဆေးခြင်း။

၉.၃ တစ်ချက်ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (One-Factor Analysis of Variance - ANOVA)

တစ်ချက် ANOVA သည် အုပ်စုသုံးခု သို့မဟုတ် သုံးခုထက်ပိုသော အုပ်စုများ၏ ပျမ်းမျှများကို နှိုင်းယှဉ်ရန် အသုံးပြုသည်။

စုစုပေါင်းနှစ်ထပ်ကိန်းပေါင်းလဒ် (Total Sum of Squares - SST):

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{..})^2$$

အုပ်စုကြားနှစ်ထပ်ကိန်းပေါင်းလဒ် (Between-Groups Sum of Squares - SSB):

$$SSB = \sum_{i=1}^{k} n_i \, (\bar{X}_{i.} - \bar{X}_{..})^2$$

အုပ်စုအတွင်းနှစ်ထပ်ကိန်းပေါင်းလဒ် (Within-Groups Sum of Squares - SSW):

$$SSW = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2$$

ANOVA ဇယား (ANOVA Table):

အရင်းအမြစ် (Source)	SS	df	MS	F
အုပ်စုကြား (Between)	SSB	k-1	MSB = SSB/(k-1)	F = MSB/MSW
အုပ်စုအတွင်း (Within)	SSW	N-k	MSW = SSW/(N-k)	
စုစုပေါင်း (Total)	SST	N-1		

ဥပမာ ၉.၃-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

မော်တော်ကားပြတင်းပေါက်များအတွက် ဆွဲအားစစ်ဆေးမှုရလဒ်များကို နှိုင်းယှဉ်ခြင်း။

၉.၄ နှစ်ချက်ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (Two-Way Analysis of Variance - ANOVA)

နှစ်ချက် ANOVA သည် အကြောင်းရင်းနှစ်ခု (Factors) ၏ သက်ရောက်မှုကို တစ်ပြိုင်နက်တည်း လေ့လာရန် အသုံးပြုသည်။

ပုံစံ (Model):

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- α_i = အကြောင်းရင်း A ၏ သက်ရောက်မှု
- β_j = အကြောင်းရင်း B ၏ သက်ရောက်မှု

- $(\alpha\beta)_{ij}$ = အပြန်အလှန်သက်ရောက်မှု (Interaction Effect)

ANOVA ဇယား (ANOVA Table):

အရင်းအမြစ် (Source)	SS	Df	MS	F
အကြောင်းရင်း A	SSA	a-1	MSA = SSA/(a-1)	F_A = MSA/MSE
အကြောင်းရင်း B	SSB	b-1	MSB = SSB/(b-1)	F_B = MSB/MSE
အပြန်အလှန် (Interaction)	SSAB	(a-1)(b-1)	MSAB = SSAB/((a-1)(b-1))	F_AB = MSAB/MSE
အမှား (Error)	SSE	ab(n-1)	MSE = SSE/(ab(n-1))	
စုစုပေါင်း (Total)	SST	abn-1		

၉.၅ စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်း (Statistical Quality Control)

စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်းသည် ထုတ်ကုန်တစ်ခု၏ အရည်အသွေးကို ထိန်းသိမ်းရန် စာရင်းအင်းနည်းလမ်းများကို အသုံးပြုသည်။

၉.၅.၁ ထိန်းချုပ်ဇယားများ (Control Charts)

$\bar{X}$ ဇယား (Chart for the Mean):

- ဗဟိုမျဉ်း (Center Line) = $\bar{\bar{X}}$
- အထက်ထိန်းချုပ်ကန့်သတ်ချက် (Upper Control Limit - UCL) = $\bar{\bar{X}} + A_2\bar{R}$
- အောက်ထိန်းချုပ်ကန့်သတ်ချက် (Lower Control Limit - LCL) = $\bar{\bar{X}} - A_2\bar{R}$

R ဇယား (Chart for the Range):

- ဗဟိုမျဉ်း (Center Line) = $\bar{R}$
- UCL = $D_4\bar{R}$

- LCL = $D_3\bar{R}$

p ဇယား (Chart for Proportion Defective):

- ဗဟိုမျဉ်း (Center Line) = $\bar{p}$
- UCL = $\bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
- LCL = $\bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

c ဇယား (Chart for Number of Defects):

- ဗဟိုမျဉ်း (Center Line) = $\bar{c}$
- UCL = $\bar{c} + 3\sqrt{\bar{c}}$
- LCL = $\bar{c} - 3\sqrt{\bar{c}}$

ဤအခန်း၏ အဓိကအချက်မှာ Chi-Square စစ်ဆေးမှုများ၊ ANOVA၊ နှင့်
စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်းတို့သည် ဒေတာခွဲခြမ်းစိတ်ဖြာခြင်းနှင့်
အရည်အသွေးမြှင့်တင်ခြင်းအတွက် အရေးပါသော ကိရိယာများဖြစ်ကြောင်း ဖော်ပြထားပါသည်။